DOCUMENT RESUME

ED 260 094 TM 850 410

AUTHOR Muraki, Eiji
TITLE Implementing Full Information Factor Analysis: TESTFACT Program.
PUB DATE Nov 84
NOTE 23p.; Paper presented at the Annual Meeting of the Psychometric Society (San Antonio, TX, November 1-3, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; *Computer Simulation; *Computer Software; Correlation; Difficulty Level; *Factor Analysis; *Guessing (Tests); *Item Analysis; Mathematical Models; Response Style (Tests); Statistical Studies; Test Items
IDENTIFIERS *Full Information Factor Analysis; Heywood Cases; *TESTFACT (Computer Program); Tetrachoric Correlation

ABSTRACT

The TESTFACT computer program and full-information factor analysis of test items were used in a computer simulation conducted to correct for the guessing effect. Full-information factor analysis also corrects for omitted items. The present version of TESTFACT handles up to five factors and 150 items. A preliminary smoothing of the tetrachoric correlation coefficient may be needed before the principal factor analysis is carried out; the new matrix is then analyzed by the MINRES method. In the simulation, stepwise full-information factor analysis of one or two factors was carried out on a 25-item data set. One model corrected for guessing and one did not. Item difficulty and chi square statistics were computed for a one-factor and a two-factor solution. Difficult items were most affected by guessing, and the guessing correction raised the tetrachoric correlation coefficients. Goodness of fit statistics were affected by additional factors or dimensionality. It was recommended that researchers using full-information factor analysis choose both the step-wise option and the guessing model in the TESTFACT program. Also the TESTFACT option which imposes constraints on item parameter estimates in maximum likelihood factor analysis should be used to avoid the Heywood case. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED260094

IMPLEMENTING FULL INFORMATION FACTOR ANALYSIS: TESTFACT PROGRAM

Eiji Muraki

National Opinion Research Center

A paper presented at the annual meeting of Psychometrica Society, 1984

TM 850 410

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

E Muraki

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TESTFACT

The TESTFACT is a highly developed and flexible FORTRAN program designed to analyze dichotomous test item responses. The program computes classical item statistics, such as item facilities and item-test correlations, and test statistics, such as coefficients of test homogeneity and test validity against an external criterion. The item and test statistics for subsets of items and examinees are also possible as an option. In addition to the classical item-test statistics, the program provides the MINRES principal factor analysis based on the item tetrachoric correlation matrix. The resulting principal factor pattern can be rotated orthogonally to the Varimax criterion (Kaiser, 1958). With the varimax solution, the pattern can be then rotated obliquely by the Promax method (Hendrickson and White, 1964). Multiple factor analysis based on tetrachoric correlations, however, have encountered difficulties because of 1) inaccuracies of calculating the coefficients as they approach +1 or -1, 2) failure to account for the effects of chance success due to multiple choice items, and 3) the lack of a statistical test for the number of factors.

Full-Information Factor Analysis

The first attempt to overcome the difficulties encountered in the factor analysis of categorical variables was made by Bock and Aitkin (1981). Their factor analysis model for dichotomcus variables is the direct extension of a unidimensional normal ogive item response model, which is given by

$$P(x_{ij}=1 \mid \underline{\theta}_i) = g_j + (1-g_j)\Phi[Z_j(\underline{\theta}_i)] \qquad (1)$$

$$= g_j + (1-g_j)\int_{-Z_j(\underline{\theta}_i)}^{\infty} \phi(t)\,dt$$

where

$$Z_j(\underline{\theta}_i) = c_j + \sum a_{jk}\theta_{ik}$$

The factor loadings a_{jk}, $k=1,2,\ldots,m$, and standard difficulty δ_j are calculated from the estimates of the slope parameter a_{jk} and intercept parameter c_j as follows:

$$\alpha_{jk} = a_{jk}/d_j,\ \delta_j = c_j/d_j,\ k=1,2,\ldots,m \qquad (2)$$

and

$$d_j^2 = 1 + \sum a_{jk}^2 \qquad (3)$$

The practical iterative estimation procedure is also developed by Bock and Aitkin (1981) based on the EM algorithm of Dempster, Laird, and Rubin (1977). Since this item factor analysis accounts for not only the pairwise joint frequencies of correct-incorrect responses, but also for additional information in higher order joint frequencies in the sample of binary scored items, the method is called full-information factor analysis (Bartholemew, 1980). Since

the model (1) contains a guessing parameter g_j for item j, the full-information method easily accomodates a correction for guessing effects. This full-information item factor analysis is implemented in the TESTFACT program. The resulting factor pattern can be rotated orthogonally and then obliquely. The TESTFACT further provides the option of a stepwise analysis to determine the significant number of factors. At each step of the full-information factor analysis, the solution from the principal factor analysis for the current number of factors is used as initial estimates of the parameters. Thus, a test of the number of factors is based on the change in marginal log likelihood as additional factors are added in the analysis. The computation of the full-information factor solution becomes exponentially heavy as the number of factors increases. The present version of the TESTFACT program can deal with up to five factors and 150 items. The scoring (the EAP estimates) of response vectors is also possible after obtaining the item estimates. Users can select the number of quadrature points for the scoring as well as the parameter estimation; otherwise a default value is set according to the number of factors extracted.

Another noteworthy feature of the full information factor analysis is that there is no necessity of assuming that all subjects are presented the same items. Missing item responses can be excluded from the analysis under the assumption

of local independence and the additive nature of log likelihood. The TESTFACT program fully utilizes this feature of the full-information factor analysis so that users can choose flexible options for the treatment of missing responses. If a user specifies OMIT=MISS, omitted responses are treated as missing responses and excluded from the likelihood equations. Otherwise, a default option, OMIT=RECODE, changes all omitted responses to incorrect ones. If the time limits of a power test are too short, an appreciable portion of examinees may not reach later items in the test. To minimize the occurrence of a spurious factor associated with item position, the program further provides a TIME option to recode all omitted items following the last item responded to as "not reached". In this option, omitted items prior to the last-responded-items are scored incorrect, and other omitted ones are treated as missings. The TESTFACT program can also accomodate the analysis of the data gathered in multiple-matrix sampling designs where each subject responds to a small subset of whole items. In a similar manner, several forms of tests linked with common items can be analyzed simultaneously. In theses options, the program can distinguish "not presented" items from "presented-but-omitted" items according to distinct response codes in a part of input data. Detailed descriptions of the EM algorithm and the full-information item facrtor analysis are presented in Bock and Aitkin (1981) and Bock (1984). Other computational options provided by the TESTFACT program

are discussed in the TESTFACT user's mannual (Wilson, Wood, and Gibbons, 1984).

The possible divergence of iterative estimation process and initial lengthy iterations of the full-information factor analysis can be avoided as much as possible if appropriate parameter values are provided. In the TESTFACT program, these starting values are immediately obtained by the principal factor analysis of the matrix of inter-item tetrachoric correlations. The matrix is corrected for missing and guessing and conditioned to be positive semi-definite so that the principal factor analysis can produce good starting values for the full-information factor analysis. In the next sections of this paper, these correction and conditioning mechanisms of the TESTFACT will be discussed. The last section of this paper presents the demonstration of guessing effects on the results of the full-information factor analysis based a simulated data set.

Correction for Guessing

Carroll (1961, p351) defines the topastic error as "the partly random, partly systematic error which results when the individual taking a multiple-choice psychological test has the opportunity to get some of his answer correct by guessing." Carroll (1961, 1983) also points out that a failure to take into account the effects of the topastic error in factor analysis, where the topastic error exists

and the items are heterogeneous in difficulty, may lead to misleading results. If the item response model with guessing parameter is used for the full-information factor analysis, the tetrachoric correlation matrix, therefore, must be corrected for guessing and then provided to the principal factor analysis to produce good starting parameter values. Carroll's (1945) correction formulas for guessing are implemented in the TESTFACT for this purpose.

Carroll's (1945) method provides a way to correct pairwise item joint frequencies (from which the tetrachoric coefficient is calculated) for effects of random correct responses with a specified probability of chance success. To present Carroll's correction method on the proportions in the 2 x 2 table for the effect of chance success, let us denote g_i and g_j are the probability of chance success for item i and j respectively. We further denote π_{ij} as the observed proportions in the original 2 x 2 table, which are affected by chance success, and π'_{ij} as the proportions in the corrected 2 x 2 table, which are unaffected by chance success. Thus, the original and corrected contingency tables are expressed in Tables 1 and 2 respectively.

Guessing parameter is the probability of passing an item by chance when the true state of mastery for the item should be classified as failure. Thus, the observed proportion of passing is the sum of the proportion of the true state of passing and the joint proportions of the corresponding guessing and the true failure state. Therefore, we obtain

TABLE 1

THE ORIGNINAL PROPORTIONS OF SUBJECTS PASSING AND FAILING ITEMS I AND J

		Item j Pass	Fail	Total
Item i	Pass	π_{11}	π_{10}	$\pi_{1.}$
	Fail	π_{01}	π_{00}	$\pi_{0.}$
	Total	$\pi_{.1}$	$\pi_{.0}$	1.0

TABLE 2

THE CORRECTED PROPORTIONS OF SUBJECTS PASSING AND FAILING ITEMS I AND J

		Item j Pass	Fail	Total
Item i	Pass	π'_{11}	π'_{10}	$\pi'_{1.}$
	Fail	π'_{01}	π'_{00}	$\pi'_{0.}$
	Total	$\pi'_{.1}$	$\pi'_{.0}$	1.0

$$\pi_{1.} = \pi'_{1.} + g_i\pi'_{0.} \tag{4}$$

$$\pi_{.1} = \pi'_{.1} + g_j\pi'_{.0} \tag{5}$$

$$\pi_{11} = \pi'_{11} + g_i\pi'_{01} + g_j\pi'_{10} + g_ig_j\pi'_{00} \tag{6}$$

and

$$\pi'_{11} + \pi'_{01} + \pi'_{10} + \pi'_{00} = 1 \tag{7}$$

From Equation (4), (5), (6), and (7), we solve the corrected proportions π's in terms of the observed proportion π and guessing parameters g_j as follows:

$$\pi'_{00} = \pi_{00}/w_iw_j \tag{8}$$

$$\pi'_{01} = (w_j\pi_{01}-g_j\pi_{00})/w_iw_j \tag{9}$$

$$\pi'_{10} = (w_i\pi_{10}-g_i\pi_{00})/w_iw_j \tag{10}$$

and

$$\pi'_{11} = 1 - \pi'_{00} - \pi'_{01} - \pi'_{10} \qquad (11)$$

where $w_i = 1 - g_i$ and $w_j = 1 - g_j$.

Conversion of the kth factor loading a_{jk} to the provisional slope estimate a_{jk} is

$$a_{jk} = a_{jk}/\sigma_j \qquad (12)$$

where $\sigma_j^2 = 1 - \sum a_{jk}^2$

The provisional intercept estimates c_j are computed from the standard difficulty δ_i and σ_j such that

$$c_j = \delta_i/\sigma_j \qquad (13)$$

since $\sigma_j = d_j^{-1}$

The standard difficulty δ_j is the normal ordinate of facility ξ_j, which is measured by the proportion of individuals passing item j. The corrected facility ξ'_j for guessing is computed by

$$\xi'_j = 1 - (1-\xi_j)/(1-g_j) \qquad (14)$$

Correction for Omitted Responses

The TESTFACT program provides three options -COMPLETE, PAIRWISE, AND RECODE- for the computation of tetrachoric correlation coefficients. If COMPLETE is chosen, only observations with no omitted responses are used. In PAIRWISE option, all observations are used, but where omitted responses occur, that item is ignored. RECODE is default, and in this option all omitted responses are recoded as incorrect responses. When the RECODE option is chosen, the observed proportions π_{10}, π_{01}, and π_{00} tend to

be inflated. Since the positive corrected proportions are obtained only if $\pi_{00}/\pi_{0.} \leq w_j$ and $\pi_{00}/\pi_{.0} \leq w_i$, negative corrected proportions are the likely result. This problem is almost always encountered because omitted responses are frequently found in cognitive testing. Since the guessing parameters are not estimated in the present version of the TESTFACT program and should be supplied as known constants in a part of input data,[1] the possible solution for this problem is to allocate omitted responses to the categories of correct and incorrect responses. This correction for omits must be made before the correction for guessing.

Let us denote n_{ij} as the observed frequency in the 3 x 3 table whose categories are pass, fail, and omit. Thus, the observed frequency table is expressed in Table 3.

If the proportions of correct and incorrect responses based on non omitted responses are denoted by p's and q's respectivley, they are computed by

$$p_i = (n_{11}+n_{10})/N.. \quad (15)$$

$$q_i = (n_{01}+n_{00})/N.. \quad (16)$$

$$p_j = (n_{11}+n_{01})/N.. \quad (17)$$

and

$$q_j = (n_{10}+n_{00})/N.. \quad (18)$$

---

[1]Guessing parameters are ordinates of the asymptote of the response function in the direction of low ability. As such, they do not depend upon the form of the response curve at higher abilities. For this reason, the guessing parameters can be satisfactorily estimated by a unidimensional item response model such as that used in the BILOG program of Mislevy and Bock (1982).

TABLE 3

OBSERVED FREQUENCIES OF SUBJECTS PASSING, FAILING, AND OMITTING ITEMS I AND J

		Item j			
		Pass	Fail	Omit	Total
Item i	Pass	n_{11}	n_{10}	n_{1x}	$n_{1.}$
	Fail	n_{01}	n_{00}	n_{0x}	$n_{0.}$
	Omit	n_{x1}	n_{x0}	n_{xx}	$n_{x.}$
	Total	$n_{.1}$	$n_{.0}$	$n_{.x}$	$n_{..}$

where $N.. = n_{11} + n_{10} + n_{01} + n_{00}$

If we can assume that omitted responses can be reallocated to correct and incorrect responses proportional to p's and q's, the following corrected frequencies n'_{ij} are obtained:

$$n'_{11} = n_{11} + p_j n_{1x} + p_i n_{x1} + p_i p_j n_{xx} \quad (19)$$

$$n'_{10} = n_{10} + q_j n_{1x} + p_i n_{x0} + p_i q_j n_{xx} \quad (20)$$

$$n'_{01} = n_{01} + p_j n_{0x} + q_i n_{x1} + q_i p_j n_{xx} \quad (21)$$

and

$$n'_{00} = n_{00} + q_j n_{0x} + q_i n_{x0} + q_i q_j n_{xx} \quad (22)$$

Marginal frequecies are computed by

$$n'_{1.} = n_{1.} + p_i n_{x.} \quad (23)$$

$$n'_{0.} = n_{0.} + q_i n_{x.} \quad (24)$$

$$n'_{.1} = n_{.1} + p_j n_{.x} \quad (25)$$

and

$$n'_{.0} = n_{.0} + q_j n_{.x} \quad (26)$$

Therefore,

$$n'_{1.} + n'_{0.} = n'_{.1} + n'_{.0} = n.. \quad (27)$$

because

$$p_i + q_i = p_j + q_j = 1 \quad (28)$$

Preliminary Smoothing of the Tetrachoric Correlation Matrix

Although the correction for omits makes the calculation of most tetrachoric correlations possible, there are still occasional instances in large size matrices where a value close to 0 appears in the minor diagonal of the tables of a few item pairwise joint frequencies. In this case, the tetrachoric correlation coefficients for these pairs are automatically set as 1. These inadmissible values are replaced by the estimated tetrachoric correlations based on Thurstone's centroid method under the assumption the matrix has one dominant common factor. This imputation method is carried out based on the assumption one dominant common factor is generally the case when cognitive abilities are measured.

The tetrachoric correlation matrix to be analyzed by the MINRES method (Harman, 1976) must be at least positive semi-definite. The corrected matrix obtained through the centroid method, on the other hand, may have negative roots. Therefore, a preliminary "smoothing" of the tetrachoric correlation coefficient matrix is needed before the principal factor analysis is carried out. The smoothed tetrachoric correlation matrix is produced from the eigen vectors associated with the positive roots, after renorming the sum of the roots to equal the number of items. The reproduced positive semi-definite tetrachoric correlation matrix is then analyzed by the MINRES method to obtain good starting values for the full-information factor analysis.

Constraints on Item Parameter Estimates

Maximum likelihood factor analysis frequenctly encounters "Heywood" case unless otherwise constrained. The TESTFACT program provides the option of imposing such a constraint on item parameter estimates by placing a prior beta distribution $\beta(p,1)$ on the unique factor loading u_j and a prior normal distribution $N(\theta_c, \sigma^2_c)$ on the intercept estimate c_j. The so-called "penalty functions" for the estimation of slope parameters to be added to the elements of the first derivative, the off-diagonal elements and the diagonal elements of the second derivatives are presented in (29), (30), and (31) respectively.

$$-2(p-1)a_{jk}/d_j^2 \qquad (29)$$

$$-4(p-1)a_{jk}a'_{jk}/d_j^4 \qquad (30)$$

and

$$2(p-1)(d_j^2-2a_{jk}^2)/d_j^4 \qquad (31)$$

where $d_j^2 = 1 + \sum a_{jk}^2$

and uniqueness $u_j^2 = d_j^{-2} = 1 - \sum a_{jk}^2$

For estimation of intercept parameters, the following terms are added to the elements of first derivative and the diagonal elements of second derivative:

$$-(c-\mu_c)/\sigma_c^2 \qquad (32)$$

and

$$1/\sigma_c^2 \qquad (33)$$

Simulation Study of Guessing Effect

The capability of incorporating guessing effects on the estimation of intercept and slope parameters is one of the advantages of the TESTFACT program. Carroll (1945, 1961, 1983) pointed out that ignoring measurement error due to a guessing effect may produce a spurious factor in factor analysis. In this last section, we domonstrate how such erroneous results have actually occurred in our simulated data.

Stepwise full-information item factor analysis of up to two factors were carried out on a simulated data set generated by the BILOG program. The model used for generating response vectors is a unidimensional three parameter normal ogive item response model. Slope parameters a_j and guessing (chance success) parameters g_j are uniformly set as 0.588 and 0.2 for all items respectively. Threshold values δ_j are sampled from a standard normal distribution. The 100 response cases to 50 items were generated. Only an even number of item responses were selected for the analysis. For each dimension of the latent trait space, ten quadrature points for one factor solution and five quadrature points for two factor solution were used. Maximum major iterations and minor iterations are set as 25 and 5 respectively. The convergence level for both major and minor iterations is set as 0.001.

Tables 4 and 5 give the estimated parameter values of the models with and without guessing constraints respectively.

TABLE 4

PARAMETER ESTIMATES OF THE MODEL WITH GUESSING

Item	Facility	One-Factor Intercept	One-Factor Slope	Two-Factor Intercept	Two-Factor Slopes	
1	0.543	-0.2006	0.5160	-0.2009	0.4813	0.3693
2	0.507	-0.3448	0.6182	-0.3389	0.5613	0.3355
3	0.459	-0.5391	0.6330	-0.5207	0.5785	0.1657
4	0.572	-0.1000	0.5483	-0.0942	0.6403	-0.2581
5	0.705	0.3876	0.5785	0.3974	0.5422	0.1896
6	0.860	1.0406	0.4889	1.2291	0.6839	-0.4617
7	0.825	0.9470	0.6951	0.9754	0.7208	0.0033
8	0.817	0.8419	0.5323	0.8660	0.5615	-0.0663
9	0.685	0.3150	0.6027	0.3295	0.5517	0.2980
10	0.790	0.7130	0.5050	0.7244	0.4922	0.1302
11	0.445	-0.5681	0.5388	-0.5817	0.5118	0.3319
12	0.536	-0.2312	0.5758	-0.2267	0.6238	-0.1374
13	0.493	-0.3850	0.5017	-0.3836	0.4657	0.2809
14	0.668	0.2556	0.6409	0.2659	0.6158	0.0518
15	0.592	-0.0247	0.6523	-0.0113	0.6178	0.2164
16	0.713	0.4235	0.5963	0.4374	0.6021	0.0589
17	0.732	0.5012	0.6080	0.5122	0.5986	0.0526
18	0.488	-0.4074	0.5296	-0.3961	0.4952	0.1358
19	0.543	-0.2067	0.5982	-0.1972	0.6238	-0.1022
20	0.714	0.4448	0.7020	0.4585	0.6790	0.1790
21	0.794	0.7287	0.4999	0.7409	0.4762	0.1781
22	0.783	0.7170	0.6192	0.7327	0.5837	0.2389
23	0.727	0.4613	0.5197	0.4753	0.5358	0.0247
24	0.696	0.3440	0.5200	0.3660	0.4584	0.4019
25	0.285	-1.4030	0.4744	-1.8699	0.8105	-0.6795

Item facilities computed from the data are presented at the second column of Table 4. Table 6 gives chi-square statistics for one factor and two factor solution.

The same analyses have been performed with prior distributions, $\beta(1.2,1)$ and $N(0,4)$. The results are presented in Tables 7 and 8 in the same manner as Tables 4 and 5.

TABLE 5

PARAMETER ESTIMATES OF THE MODEL WITHOUT GUESSING

Item	One-Factor Intercept	One-Factor Slope	Two-Factor Intercept	Two-Factor Slope	
1	0.1178	0.4011	0.1217	0.3562	0.2395
2	0.0213	0.4493	0.0251	0.3892	0.4238
3	-0.1098	0.4252	-0.1073	0.3955	0.1624
4	0.1987	0.4164	0.2181	0.5360	-0.2647
5	0.6025	0.4888	0.6064	0.4436	0.2055
6	1.1851	0.4497	1.3211	0.6148	-0.3184
7	1.1054	0.6315	1.1164	0.6361	0.0789
8	1.0032	0.4786	1.0067	0.4605	0.1250
9	0.5412	0.4984	0.5496	0.4463	0.2687
10	0.8872	0.4574	0.8938	0.4645	0.0339
11	-0.1472	0.3855	-0.1448	0.3463	0.1884
12	0.1005	0.4295	0.1065	0.4743	-0.0978
13	-0.0167	0.3514	-0.0147	0.3124	0.2125
14	0.4935	0.5241	0.4964	0.5141	0.0740
15	0.2650	0.5291	0.2692	0.5150	0.1456
16	0.6325	0.5098	0.6397	0.5164	0.0562
17	0.7028	0.5353	0.7080	0.5328	0.0724
18	-0.0301	0.3647	-0.0281	0.3255	0.2119
19	0.1206	0.4529	0.1259	0.4856	-0.0541
20	0.6556	0.5715	0.6589	0.5299	0.2017
21	0.9025	0.4574	0.9113	0.4223	0.2123
22	0.8939	0.5476	0.9095	0.4900	0.3060
23	0.6624	0.4436	0.6704	0.4583	0.0244
24	0.5613	0.4322	0.5643	0.3863	0.1947
25	-0.5749	0.1587	-0.5742	0.1630	0.0191

TABLE 6

CHI-SQUARE STATISTICS FOR THE TWO-FACTOR STEPWISE ANALYSIS WITH AND WITHOUT GUESSING

	No Guessing Chi-Square	D.F.	P	Guessing Chi-Square	D.F.	P
One-Factor	15431.78	930	0.000	15423.07	930	0.000
Two-Factor	15386.81	906	0.000	15383.01	906	0.000
Change	44.97	24	0.006	40.06	24	0.021

Drastic changes appear in the intercept estimates between

TABLE 7

CONSTRAINED PARAMETER ESTIMATES OF THE MODEL WITH GUESSING

Item	One-Factor Intercept	One-Factor Slope	Two-Factor Intercept	Two-Factor Slopes	
1	-0.2006	0.5149	-0.1999	0.4703	0.3631
2	-0.3444	0.6163	-0.3421	0.5536	0.3827
3	-0.5382	0.6305	-0.5212	0.5774	0.1754
4	-0.1000	0.5467	-0.0957	0.6538	-0.2557
5	0.3871	0.5769	0.3964	0.5393	0.1986
6	1.0391	0.4870	1.2291	0.6987	-0.4488
7	0.9453	0.6925	0.9719	0.7181	0.0284
8	0.8408	0.5306	0.8631	0.5608	-0.0498
9	0.3145	0.6011	0.3285	0.5469	0.3058
10	0.7122	0.5037	0.7228	0.4910	0.1306
11	-0.5674	0.5370	-0.5791	0.5031	0.3305
12	-0.2310	0.5740	-0.2276	0.6277	-0.1278
13	-0.3847	0.5003	-0.3838	0.4598	0.2920
14	0.2551	0.6393	0.2648	0.6151	0.0645
15	-0.0248	0.6507	-0.0122	0.6130	0.2373
16	0.4229	0.5947	0.4361	0.6013	0.0711
17	0.5005	0.6064	0.5111	0.5990	0.0629
18	-0.4070	0.5279	-0.3961	0.4906	0.1504
19	-0.2066	0.5966	-0.1991	0.6355	-0.1067
20	0.4440	0.7000	0.4570	0.6723	0.2040
21	0.7279	0.4986	0.7392	0.4723	0.1841
22	0.7160	0.6175	0.7318	0.5758	0.2611
23	0.4607	0.5185	0.4739	0.5354	0.0314
24	0.3435	0.5188	0.3662	0.4517	0.4167
25	-1.3904	0.4626	-1.6089	0.6639	-0.4167

the models with and without guessing, as shown in Tables 4 and 5. The estimates of the harder items are more affected by the guessing effect than the estimates of the easy items. Just like item facilities, the intercept parameters in the model (1) seems to be very sensitive to the guessing correction. On the other hand, a consistent pattern appears in the change of slope estimates between the models with and without guessing. The first slope estimates of the model

TABLE 8

CONSTRAINED PARAMETER ESTIMATES OF THE MODEL WITHOUT GUESSING

Item	One-Factor Intercept	One-Factor Slope	Two-Factor Intercept	Two-Factor Slopes	
1	0.1173	0.4005	0.1212	0.3571	0.2362
2	0.0207	0.4486	0.0245	0.3906	0.4218
3	-0.1102	0.4246	-0.1078	0.3954	0.1597
4	0.1981	0.4157	0.2176	0.5356	-0.2711
5	0.6015	0.4878	0.6053	0.4432	0.2038
6	1.1833	0.4481	1.3101	0.6033	-0.3067
7	1.1032	0.6292	1.1139	0.6335	0.0766
8	1.0017	0.4772	1.0051	0.4583	0.1267
9	0.5403	0.4975	0.5486	0.4459	0.2680
10	0.8860	0.4562	0.8927	0.4641	0.0284
11	-0.1476	0.3849	-0.1452	0.3460	0.1881
12	0.1000	0.4288	0.1057	0.4720	-0.0985
13	-0.0171	0.3509	-0.0152	0.3131	0.2117
14	0.4925	0.5231	0.4954	0.5141	0.0683
15	0.2642	0.5282	0.2683	0.5151	0.1413
16	0.6315	0.5088	0.6386	0.5159	0.0514
17	0.7016	0.5342	0.7067	0.5317	0.0681
18	-0.0305	0.3642	-0.0286	0.3257	0.2110
19	0.1200	0.4521	0.1252	0.4846	-0.0573
20	0.6544	0.5702	0.6574	0.5298	0.1955
21	0.9013	0.4562	0.9101	0.4217	0.2116
22	0.8924	0.5461	0.9083	0.4896	0.3053
23	0.6614	0.4426	0.6695	0.4581	0.0199
24	0.5604	0.4314	0.5629	0.3875	0.1851
25	-0.5748	0.1584	-0.5741	0.1628	0.0179

with guessing are higher over whole items than the estimates of the model without guessing. The higher estimated values are resulted by excluding the topastic error from the slope parameter estimation. This corresponds to the phenomenon that the guessing correction raises the tetrachoric correlation coefficients. In our simulated data set, the average of the corrected coefficients is 0.240 while the average

of the uncorrected ones is 0.163. If the responses contains the topastic variance, the introduction of guessing parameter into the multidimensional item response model makes communality increase, and consequently may facilitate the identification of dominant factor pattern.

As shown in Table 6, for the model with guessing parameter, the reduction of chi-square fit statistics caused by adding additional factor is significant at 0.01 probability level while the significant reduction does not occur for the guessing model, which is compatible with the generating model. This result is not dramatic but highly suggestive of the guessing effect on the factor analysis. As Carroll (1945, 1961, 1983) pointed out for the case of tetrachoric factor analysis, the full-information factor model which fails to take account the guessing effect also seems to have produced a spurious second factor when the response variables are affected by topastic error and item difficulties are heterogeneous. The values of the second slope estimates vary considerably. However, the absolute values are never more than 0.43. This may also support that the second slope estimates are spurious.

Adding additonal dimensionality to the model certainly decrease the fit statitics. The difference of the chi-square statistics between the guessing and non-guessing one-factor models is 8.71, but the difference of the statistics between the guessing and non-guessing models becomes

smaller, 3.80, in the two-factor case. This seems to suggest that the topastic error in the generated data have been absorbed by the additional dimensionality of the non-guessing model. The relationship between the guessing effect and the slope estimates needs to be investigated further. To avoid the risk of a fault interpretation of spurious factors, researchers use the full-information factor analysis for exploration purpose are recommended to choose both the step-wise option and the guessing model in the TESTFACT program.

Item 25 is the most difficult item, and this item is affected the most by the guessing correction. Through the iteration process of the estimation, both the intercept and slope estimates of this item becomes larger and larger. With imposing constraints on the estimation, this phenomenon becomes less severe. Our mild constraints also can maintain the relative magnitude of estimated values between the analyses with and without prior distributions. Whenever the "Heywood" case is suspected in the first analysis, the prior option of the TESTFACT program is recommended.

BIBLIOGRAPHY

Bartholomew, D. J. Factor analysis for categorical data. Journal of the Royal Statistical Society, Series B, 1980, 42, 293-321.

Bock, R. D. (1984) (This symposium)

Bock, R. D., Aitkin, M. Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. Psychometrika, 1981, 46, 443-459.

Bock, R. D., and Mislevy, R. J. BILOG: Maximum Likelihood Item Analysis and Test Scoring Binary Logistic Models. Mooresville, Indiana: Scientific Software, Inc., 1984.

Carroll, J. B. The effect of difficulty and chance success on correlations between items and between tests. Psychometrika, 1945, 10, 1-19.

Carroll, J. B. The nature of the data, or how to choose a correlation coefficent. Psychometrika, 1961, 26, 347-372.

Carroll, J. B. The difficulty of a test and its factor composition revisited. In Wainer, H. & Messick, S. (Eds.), Principals of Modern Psychological Measurement, New Jersey: Lawrence Erlbaum Associates, Publishers, 1983, 257-282.

Dempster, A. P., Laird, N., and Rubin D. B. Maximum likelihood from incomplete data via the EM algorithm (with Discussion). Journal of the Royal Statistical Society, Series B, 1977, 39, 1-38.

Harman, H. H Modern Factor Analysis. Chicago: The University of Chicago Press, 1976.

Hendrickson, A. E., and White, P. O. PROMAX: A quick method for rotation to oblique simple structure. British Journal of Mathematical and Statistical Psychology, 1964, 17, 65-70.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 287-200.

BEST COPY

Mislevy, R. J., and Bock, R. D. Implementation of the EM Algorithm in the Estimation of Item Parameters: The BILOG Computer Program. A paper presented at the 1982 IRT/CAT Invitational Conference at the University of Minnesota, July 20-30.

Wilson, D, Wood, R., and Gibbons, R. TESTFACT:Test Scoring and Item Factor Analysis. Mooresville, Indiana: Scientific Software, Inc., 1984.